STRONG ORDER OF CONVERGENCE OF A FULLY DISCRETE APPROXIMATION OF A LINEAR STOCHASTIC VOLTERRA TYPE EVOLUTION EQUATION

MIHÁLY KOVÁCS AND JACQUES PRINTEMS

Abstract. In this paper we investigate a discrete approximation in time and in space of a Hilbert space valued stochastic process $\{u(t)\}_{t\in[0,T]}$ satisfying a stochastic linear evolution equation with a positive-type memory term driven by an additive Gaussian noise. The equation can be written in an abstract form as

$$\mathrm{d}u + \left(\int_0^t b(t-s)Au(s)\,\mathrm{d}s\right)\mathrm{d}t = \mathrm{d}W^Q,\ t\in(0,T];\quad u(0)=u_0\in H,$$

where W^Q is a Q-Wiener process on $H=L^2(\mathcal{D})$ and where the main example of b we consider is given by

$$b(t)=t^{\beta-1}/\Gamma(\beta),\quad 0<\beta<1.$$

We let A be an unbounded linear self-adjoint positive operator on H and we further assume that there exist $\alpha>0$ such that $A^{-\alpha}$ has finite trace and that Q is bounded from H into $D(A^\kappa)$ for some real κ with $\alpha-\frac{1}{\beta+1}<\kappa\le\alpha$.

The discretization is achieved via an implicit Euler scheme and a Laplace transform convolution quadrature in time (parameter $\Delta t=T/n$), and a standard continuous finite element method in space (parameter h). Let $u_{n,h}$ be the discrete solution at $T=n\Delta t$. We show that

$$\left(\mathbb{E}\|u_{n,h}-u(T)\|^2\right)^{1/2}=\mathcal{O}(h^\nu+\Delta t^\gamma),$$

for any $\gamma<(1-(\beta+1)(\alpha-\kappa))/2$ and $\nu\le\frac{1}{\beta+1}-\alpha+\kappa$.

1. Introduction

Let $\mathcal{D}$ be a bounded domain in $\mathbb{R}^d$, $d\in\mathbb{N}$, and let u be a real-valued stochastic process solution of the equation formally written as

$$\frac{\partial u(x,t)}{\partial t}-\int_0^t b(t-s)\Delta u(x,s)\,\mathrm{d}s=\dot{\xi}(x,t),\quad (x,t)\in\mathcal{D}\times(0,T],\tag{1.1}$$

together with the initial condition $u(x,0)=u_0(x)$, $x\in\mathcal{D}$, and boundary condition $u|_{\partial\mathcal{D}}=0$. Here, $\dot{\xi}$ is a zero mean real valued Gaussian noise and the time kernel b is assumed to be real-valued and of positive type; i.e., that for any $T>0$, the kernel b belongs to $L^1(0,T)$ and for any continuous function f on $[0,T]$ the following inequality holds:

$$\int_0^T\int_0^t b(t-s)f(s)f(t)\,\mathrm{d}s\,\mathrm{d}t\ge 0.\tag{1.2}$$

2000 *Mathematics Subject Classification.* 34A08, 45D05, 60H15, 60H35, 65M12, 65M60.

Key words and phrases. Stochastic Volterra equation, fractional differential equation, finite elements method, convolution quadrature, Euler scheme, strong order.

The deterministic version of such problems can be used to model viscoelasticity or heat conduction in materials with memory (see [13] for references). When b is smooth, these equations exhibit a hyperbolic behaviour, whereas if b has a weak singularity at $t = 0$ (for example a Riesz potential), they exhibit certain parabolic features. In particular, when

$$b(t) = t^{\beta-1}/\Gamma(\beta), \quad 0 < \beta < 1, \tag{1.3}$$

the homogeneous deterministic equation has a smoothing property which correspond to the inequality

$$\|u^{(m)}(t)\|_{H^{2r}(\mathbb{R})} \leq C\, t^{-(\beta+1)r-m}\|u_0\|_{L^2(\mathcal{D})}, \tag{1.4}$$

where $|r| \leq 1$ if $m \geq 1$ and where $0 \leq r \leq 1$ if $m = 0$, but with no further smoothing in the spacial variables (see e.g. [13, Theorem 5.5]). The framework of this paper allows for slightly more general kernels but with similar smoothing effects and, in particular, they are of positive type. Hence, together with the positivity of the operator $-\Delta$, the deterministic equation will remain parabolic in character.

Next we introduce an abstract framework to describe the noise and equation (1.1) more precisely. Let Q be a self-adjoint, nonnegative linear operator on $H := L^2(\mathcal{D})$ and W^Q be a Wiener process in H with covariance operator Q (or, simply, Q-Wiener process). We set $A = -\Delta$, $D(A) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$ and $H = L^2(\mathcal{D})$. Then A can be seen as an unbounded linear operator on H with domain $D(A)$. For b given in (1.3) and under reasonable assumptions on ∂D, our main assumption concerning Q in (1.1) is that $A^\kappa Q$ defines a bounded operator on $L^2(\mathcal{D})$ with $d/2 - 1/(\beta + 1) < \kappa < d/2$.

If we write $u(t) = u(\cdot, t)$, considered as a H-valued stochastic process, then (1.1) can be rewritten in the abstract Itô form as

$$\mathrm{d}u(t) + \left(\int_0^t b(t-s)Au(s)\,\mathrm{d}s\right)\mathrm{d}t = \mathrm{d}W^Q(t), \quad t \in (0, T], \tag{1.5}$$

with initial condition $u(0) = u_0 \in H$.

While the literature on numerical methods for deterministic infinite dimensional Volterra equations is abundant (see, for example, [1, 6, 12, 13, 20], which is a very incomplete list), the numerical analysis of stochastic Volterra equations is completely missing. We are only aware of [7] where an algorithm is described and numerical experiments are performed with no error analysis given. We will consider a numerical approximation of (1.5) by means of an Euler scheme and a Laplace transform convolution quadrature in time together with a finite element method in space. Let $n \geq 1$ an integer, $\Delta t = T/n$ and $t_k = k\,\Delta t$, $k = 0, \ldots, n$. Let also $\{V_h\}_{h>0}$ be a family of finite dimensional subspaces of $D(A^{1/2}) = H_0^1(\mathcal{D})$. For each $1 \leq k \leq n$, we seek for an approximation of $u(t_k)$ in V_h by $u_{k,h}$ defined by the following induction:

$$(u_{k,h} - u_{k-1,h}, v_h) + \Delta t \sum_{j=1}^{k} \omega_{k-j}(Au_{j,h}, v_h) = \sqrt{\Delta t}(Q^{1/2}\chi_k, v_h), \quad k \geq 1, \tag{1.6}$$

for any $v_h \in V_h$, where $\sqrt{\Delta t}\,\chi_k$ is the noise increment and where $(\cdot,\cdot)$ is the inner product of H. The approximation of the convolution term in (1.5) is achieved via

a quadrature rule such that for any continuous function f on $[0,T]$,

$$\sum_{j=1}^{k} \omega_{k-j} f(t_j) \sim \int_0^{t_k} b(t_k - s) f(s) \,\mathrm{d}s = (b \star f)(t_k).$$

Then, the approximation of $b\star f$ on the time grid t_k, $k = 0, \dots, n$, is obtained from a discrete convolution with the values of f on the same grid. Before going into details, let us point out that not any quadrature rule can be chosen. In particular, it will be important for the chosen quadrature to satisfy a discrete analogue of (1.2).

In order to understand the specific quadrature rule used in this paper, we will take the example of the Riesz kernel (1.3). Let us note that in this case the Laplace transform of b is $z^{-\beta}$ and the term $b \star \Delta u$ in (1.1) can be seen as the fractional integral $(\partial/\partial t)^{-\beta}(\Delta u)$. Then, the idea is to use the same Euler approximation of $\partial/\partial t$ in both terms on the left hand side of (1.1). Since the discrete Laplace transform of the implicit Euler scheme is $(1-z)/\Delta t$, one chooses the quadrature weights to have discrete Laplace transform $((1-z)/\Delta t)^{-\beta}$.

Such a convolution quadrature has been introduced in [9, 10]. It was motivated by the fact that the main properties of the solution of the homogeneous problem, like stability, existence, or regularity, are largely determined by the distribution of the frequencies of the kernel (by means of its Fourier or Laplace transform), especially when the kernel has weak singularities or when it exhibits different behaviour at different time scales. Since, by construction, the discrete Laplace transform of the quadrature kernel is closely related to the Laplace transform of the continuous kernel, it is thus possible to carry over frequency domain conditions from the continuous problem to the discretization and thereby obtain stable approximations. Moreover, this kind of quadrature rule inherits the rate of approximation from the time integrator of $\partial/\partial t$. In the context of stochastic PDEs, we think that it is important to make sure that the deterministic part of the scheme is stable and that the perturbations are due to the noise only.

Although the analysis in the present paper allows for kernels slightly more general than (1.3), we follow the same idea: the convolution quadrature weights $\{\omega_k\}$ in (1.6) will be defined by means of the Laplace transform of the kernel b. Therefore, we choose the quadrature coefficients to have generating function $\widehat{b}((1-z)/\Delta t)$ where $\widehat{b}$ denotes the Laplace transform of b; that is,

$$\sum_{n=0}^{+\infty} \omega_k z^k = \widehat{b}\left(\frac{1-z}{\Delta t}\right), \quad |z| < 1. \tag{1.7}$$

We will not focus here on practical algorithms for the computations of the quadrature weights and we refer the reader to, for example, [10].

While precise conditions on the kernel b are postponed to Sections 2 and 4, we can already state our main result, Theorem 5.1, with the above notations in the case of the specific kernel (1.3) when $\mathcal{D}$ is a convex polygonal domain using continuous, piecewise linear finite elements. We shall prove a (strong) error estimate of the form

$$\left(\mathbb{E}(\|u_{n,h} - u(T)\|^2)\right)^{1/2} \le C(\Delta t^\gamma + h^\nu),$$

where $\gamma < (1 - (\beta+1)(d/2 - \kappa))/2$ and $\nu < 1/(\beta+1) - d/2 + \kappa$. Let us note that we recover the known order of convergence for the heat equation (see [8, 16, 21]) when $\beta \to 0$.

The paper is organized as follows. In Section 2 we introduce notations, recall some basic preliminary results, and state our main assumptions on A, Q and b. We note that Assumptions (2.8)–(2.9) on A and Q could be replaced by a single, somewhat sharper, assumption as discussed in Remarks 2.8, 3.5, 4.7 and 5.2. It is, however, harder to check in most cases. In Section 3 we study the space semi-discretization of (1.1) and strong error estimates are derived for smooth initial data under minimal regularity assumptions (Assumption 1) on b. In Section 4 we prove strong error estimates for the time semi-discrete scheme with non-smooth initial data. One of the key results in this direction is Theorem 4.1, where we prove a general l^p-stability result on Lubich's convolution quadrature based on the Backward Euler method for deterministic Volterra equations. Interestingly, this stability result implies (Corollary 4.2) that the time-discrete scheme exhibits the same smoothing effect in time as the solution under Assumption 1 on b. However, in order to obtain optimal convergence rates for the stochastic problem we need to put a further regularity restriction on b in Subsection 4.2, Assumption 2, which is in fact common in the deterministic literature for nonsmooth initial data. Indeed, Assumption 2 implies that the deterministic equation has an analytic resolvent family while Assumption 1 only implies that the deterministic equation is parabolic. Unlike for equations with no memory term, these two notions are not equivalent (See [17, Chapter 1, Section 3]). As far as we know the derivation of nonsmooth initial data estimates using only parabolicity (Assumption 1) remains an open problem. Finally, in the last section, we gather the results from the preceding sections and consider the fully discrete scheme.

2. Notations and preliminairies

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two Banach spaces and let $\mathcal{B}(X,Y)$ denote the space of bounded linear operators from X into Y endowed with the norm $\|B\|_{\mathcal{B}(X,Y)} = \sup_{x\in X} \|Bx\|_Y/\|x\|_X$. When $X = Y$, we use the shorter notation $\mathcal{B}(X)$ for $\mathcal{B}(X,X)$. If X is a Banach space and I is an interval in $\mathbb{R}$ then, $L^p(I,X)$, $1 \le p < \infty$, denotes the space of functions $f : I \to X$ which are measurable and $t \to \|f(t)\|^p$ is integrable on I, equipped with the usual norm. If $p = \infty$ then $L^\infty(I,X)$, denotes the space of functions $f : I \to X$ which are measurable and $t \to \|f(t)\|$ is essentially bounded on I endowed with the usual norm.

Throughout this paper, H denotes a real separable Hilbert space with inner product $(\cdot,\cdot)$ and associated norm $\|\cdot\|$. We consider the stochastic Volterra equation given in the abstract Itô form as

$$\mathrm{d}u + \left(\int_0^t b(t-s)Au(s)\,\mathrm{d}s\right)\mathrm{d}t = \mathrm{d}W^Q, \quad t \in (0,T]; \quad u(0) = u_0 \in H, \tag{2.1}$$

where the process $\{u(t)\}_{t\in[0,T]}$ is a H-valued stochastic process, A is a densely defined, nonnegative self-adjoint unbounded operator on H with compact inverse, and W^Q is a Q-Wiener process in H on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The weak solution of (2.1) is a mean-square continuous H-valued process satisfying

$$(u(t),\eta) + \int_0^t \int_0^r b(r-s)\,(u(s), A^*\eta)\,\mathrm{d}s\,\mathrm{d}r = (u_0,\eta) + \int_0^t \left(\eta,\, \mathrm{d}W^Q(s)\right),$$

for all $\eta \in D(A^*)$ almost surely for all $t \in [0,T]$.

It is well known that such assumptions on A implies the existence of a sequence of nondecreasing positive real numbers $\{\lambda_k\}_{k\ge1}$ and an orthonormal basis $\{e_k\}_{k\ge1}$

of H such that

$$Ae_k = \lambda_k e_k, \quad \lim_{k\to+\infty} \lambda_k = +\infty. \tag{2.2}$$

We define classically, by means of the spectral decomposition of A, the domains $D(A^s)$ of fractional powers $s \in \mathbb{R}$ of A and we set

$$\|v\|_s = \|A^{s/2}v\|, \quad v \in D(A^{s/2}).$$

Remark 2.1. We note that since A is nonnegative self-adjoint, $-A$ generates an analytic contraction semigroup on H. Moreover, for any $\theta < \pi$, there exists $M_\theta \geq 1$ such that the following holds:

$$\|(zI + A)^{-1}\|_{\mathcal{B}(H)} \leq \frac{M_\theta}{|z|}, \quad \text{for any } \ z \in \Sigma_\theta,$$

where $\Sigma_\theta = \{z \in \mathbb{C}\backslash\{0\},\ |\arg(z)| < \theta\}$.

Let $\mathcal{L}_1(H)$ denote the set of nuclear operators from H to H; that is, $T \in \mathcal{L}_1(H)$ if there are sequences $\{a_j\}, \{b_j\} \subset H$ with $\sum_{j=1}^\infty \|a_j\|\|b_j\| < \infty$ and such that

$$Tx = \sum_{j=1}^{\infty} (x, b_j)a_j, \quad x \in H.$$

Sometimes these operators are referred to as trace class operators. For $T \in \mathcal{L}_1(H)$ we define $\mathrm{Tr}(T)$, the trace of T, by

$$\mathrm{Tr}(T) = \sum_{n=1}^{+\infty} (Be_n, e_n),$$

where $\{e_n\}$ is an orthonormal basis of H. This definition turns out to be independent of the choice basis. Furthermore, if $L \in \mathcal{L}_1(H)$ and $M \in \mathcal{B}(H)$, then $LM, ML \in \mathcal{L}_1(H)$ and

$$\mathrm{Tr}(LM) = \mathrm{Tr}(ML). \tag{2.3}$$

If L is also symmetric and nonnegative, then

$$\mathrm{Tr}(LM) \leq \mathrm{Tr}(L)\|M\|_{\mathcal{B}(H)}. \tag{2.4}$$

Hilbert-Schmidt operators play also an important role in this paper. An operator $L \in \mathcal{B}(H)$ is Hilbert-Schmidt if $L^*L \in \mathcal{L}_1(H)$ or, equivalently, $LL^* \in \mathcal{L}_1(H)$. We denote by $\mathcal{L}_2(H)$ the space of such operators. It is a Hilbert space under the norm

$$\|L\|_{\mathcal{L}_2(H)} = \left(\mathrm{Tr}(L^*L)\right)^{1/2} = \left(\mathrm{Tr}(LL^*)\right)^{1/2}. \tag{2.5}$$

Our analysis will also use the Laplace transform. Let $f : \mathbb{R}_+ \to H$ be subexponential; i.e., that for any $\varepsilon > 0$ the function $t \mapsto f(t)e^{-\varepsilon t}$ belongs to $L^1(\mathbb{R}_+, H)$. We define the Laplace transform of $\widehat{f} : \mathbb{C}_+ \to H$ by

$$\widehat{f}(z) = \int_0^{+\infty} f(t)e^{-zt}\, dt, \quad \mathrm{Re}\, z > 0,$$

where we have used the same notation H for the complexification of H. We denote by $\star$ the Laplace convolution product on $[0, t]$ of two locally integrable subexponential functions $f, g \in L^1_{loc}(\mathbb{R}_+, H)$ defined as

$$(f \star g)(t) = \int_0^t f(t - s)g(s)\, \mathrm{d}s.$$

It is well known that $f \star g \in L^1_{loc}(\mathbb{R}_+, H)$ is subexponential and

$$\widehat{f \star g}\,(z) = \widehat{f}(z)\,\widehat{g}(z), \quad \operatorname{Re} z > 0.$$

2.1. **Main assumptions.** Next we state the main assumptions concerning the kernel b and the operators A and Q, which will be used throughout this paper.

Regarding b, first note that property (1.2) can be characterized by means of the Laplace transform $\widehat{b}$ of b. It is equivalent to say that $\operatorname{Re}(\widehat{b}(\lambda)) \geq 0$ for any $\operatorname{Re}\lambda > 0$ (see [15] or [17, page 38]). Now it is clear that the positivity property (1.2) is not sufficient, in general, to ensure smoothing effects like (1.4) when working with kernels that are more general than (1.3). This is why, following [3] and [14], we will impose stronger conditions on b.

Assumption 1. *The kernel $0 \neq b \in L^1_{loc}(\mathbb{R}_+)$, is 3-monotone; that is, b, $-\dot{b}$ are nonnegative, nonincreasing, convex, and $\lim_{t\to\infty} b(t) = 0$. Furthermore,*

$$\rho := 1 + \frac{2}{\pi}\sup\{|\arg\widehat{b}(\lambda)|,\ \operatorname{Re}\lambda > 0\} \in (1,2). \tag{2.6}$$

In the special case of the Riesz kernel given in (1.3) one can easily show that $\rho = 1 + \beta$. From now on we set $\beta = \rho - 1$ with ρ defined by (2.6).

Remark 2.2. It follows from [17, Proposition 3.10] that for 3-monotone and locally integrable kernels b, condition (2.6) is equivalent to

$$\lim_{t\to 0} \frac{\frac{1}{t}\int_0^t sb(s)\,ds}{\int_0^t -s\dot{b}(s)\,ds} < +\infty. \tag{2.7}$$

Also note that, by (2.6), we have that $\operatorname{Re}(\widehat{b}(\lambda)) \geq 0$ for $\operatorname{Re}\lambda > 0$ and hence b satisfies (1.2).

For A and Q we suppose that there exists numbers $\alpha > 0$ and $\kappa \in \mathbb{R}$ such that

$$\operatorname{Tr}(A^{-\alpha}) < +\infty, \tag{2.8}$$

$$A^\kappa Q \in \mathcal{B}(H), \quad \alpha - \frac{1}{\rho} < \kappa \leq \alpha. \tag{2.9}$$

2.2. **The nonhomogeneous deterministic problem.** Given $f \in L^1([0,T];H)$, Assumption 1 together with the fact that A is positive and self-adjoint implies that the deterministic problem,

$$\dot{u}(t) + \int_0^t b(t-s)Au(s)\,ds = f(t), \quad t \in (0,T], \quad u(0) = u_0 \in H, \tag{2.10}$$

is well posed for all $T > 0$. Indeed, there exists a resolvent family $\{S(t)\}_{t\geq 0} \subset \mathcal{B}(H)$ which is strongly continuous for $t \geq 0$, differentiable for $t > 0$ and uniformly bounded by 1, see [17, Corollary 1.2 and Corollary 3.3]. The unique mild solution of (2.10) is given by the following variation of parameter formula [17, Proposition 1.2]

$$u(t) = S(t)u_0 + \int_0^t S(t-s)f(s)\,ds,\ t \in [0,T].$$

Remark 2.3. The positivity of the kernel b defined in (1.2), together with the positivity of the operator A already allows for the construction of a unique solution to (2.10) using an energy argument, see [17, Corollary 1.2]. Assumption 1 gives further integrability and smoothing properties for $\{S(t)\}_{t\geq 0}$.

Note that such a resolvent family does not satisfy the semi-group property because of the non local feature of the memory term in (2.10). Nevertheless, it can be written explicitly using the spectral decomposition (2.2) of A as

$$S(t)v = \sum_{k=1}^{+\infty} s_k(t)(v, e_k)e_k, \tag{2.11}$$

where the functions $s_k(t)$ are the solutions of the ordinary differential equations

$$\dot{s_k}(t) + \lambda_k \int_0^t b(t-s)s_k(s)\, ds = 0, \quad s_k(0) = 1. \tag{2.12}$$

The next proposition summarizes the main properties of the functions $\{s_k\}_{k\geq 1}$.

Proposition 2.4. *Suppose that b satisfies Assumption 1 and let $\rho \in (1,2)$ as defined in (2.6). Then $\lim_{r\to\infty} s_k(r) = 0$ for all $k \geq 1$ and there exists $C_0 > 0$ such that for any $k \geq 1$,*

$$\|s_k\|_{L^\infty(\mathbb{R}_+)} \leq 1, \tag{2.13}$$

$$\|\dot{s_k}\|_{L^1(\mathbb{R}_+)} \leq C_0, \tag{2.14}$$

$$\|t\dot{s_k}\|_{L^1(\mathbb{R}_+)} \leq C_0\, \lambda_k^{-1/\rho}, \tag{2.15}$$

$$\|s_k\|_{L^1(\mathbb{R}_+)} \leq C_0\, \lambda_k^{-1/\rho}. \tag{2.16}$$

Proof. Estimate (2.13) follows from [17, Corollary 1.2], inequalities (2.14) and (2.15) can be found in [14, Proposition 6] and estimate (2.16) is shown in [3, Lemma 3.1] where also the fact $\lim_{r\to\infty} s_k(r) = 0$ for all $k \geq 1$ is shown in the proof of the lemma. □

Smoothing effects of the resolvent family $\{S(t)\}_{t\geq 0}$ when b satisfies Assumption 1 can be now easily proved using Proposition 2.4.

Proposition 2.5. *Let b and ρ as in Proposition 2.4. Then for any $t > 0$, there exist a constants $C_0, C_1 > 0$ such that for any $0 \leq s \leq 2/\rho$ and $0 \leq s' \leq 2$,*

$$\|A^{s/2}S(t)\|_{\mathcal{B}(H)} \leq C_0\, t^{-s\rho/2}, \quad t > 0, \tag{2.17}$$

$$\|A^{-s'/2}\dot{S}(t)\|_{\mathcal{B}(H)} \leq C_1 \|b\|_{L^1(0,t)}^{s'/2}\, t^{s'/2-1}, \quad t > 0. \tag{2.18}$$

Proof. For any $\delta \in (0,1)$ and any $k \geq 1$, Hölder's inequality, (2.14) and (2.15) yields

$$\begin{aligned}
\int_0^{+\infty} u^\delta |\dot{s_k}(u)|\, \mathrm{d}u &= \int_0^{+\infty} u^\delta |\dot{s_k}(u)|^\delta |\dot{s_k}(u)|^{1-\delta}\, \mathrm{d}u \\
&\leq \left(\int_0^{+\infty} u|\dot{s_k}(u)|\, \mathrm{d}u\right)^\delta \left(\int_0^{+\infty} |\dot{s_k}(u)|\, \mathrm{d}u\right)^{1-\delta} \\
&\leq C_0\, \lambda_k^{-\delta/\rho}.
\end{aligned}$$

Note, that the previous final estimate also holds for $\delta = 0, 1$ by (2.14) and (2.15). Then, since $s_k(t) = -\int_t^{+\infty} u^{-\delta}u^\delta \dot{s_k}(u)\, du$ as $\lim_{r\to\infty} s_k(r) = 0$ for all $k \geq 1$ by Proposition 2.4, we can conclude that

$$|s_k(t)| \leq C_0\, t^{-\delta}\lambda_k^{-\delta/\rho}, \quad t > 0, \quad \delta \in [0,1]. \tag{2.19}$$

Thus, for any $s \in [0, 2/\rho]$ and $x \in H$, (2.19) with $0 \le \delta = \rho s/2 \le 1$ implies

$$\|A^{s/2} S(t)x\|^2 = \sum_{k\ge 1} \lambda_k^s \, s_k(t)^2 (x, e_k)^2 \le C_0 \, t^{-\rho s/2} \|x\|^2,$$

which is (2.17). To show (2.18), we use [17, Corollary 3.3] which states that under Assumption 1 and since 0 belongs to the resolvent set of A, there is $M > 0$ such that

$$\|\dot{S}(t)x\| \le M t^{-1} \|x\|, \quad x \in H, \quad t > 0. \tag{2.20}$$

On the other hand, we can bound $\dot{S}(t)x$ for $x \in D(A)$ as follows:

$$\|\dot{S}(t)x\|^2 = \sum_{k\ge 1} (\dot{s_k}(t))^2 (x, e_k)^2 \tag{2.21}$$

$$= \sum_{k\ge 1} \lambda_k^2 \left(\int_0^t b(t-s) s_k(s) ds \right)^2 (x, e_k)^2 \le \|b\|_{L^1(0,t)}^2 \|Ax\|^2, \tag{2.22}$$

where we have used (2.12) and (2.13). Finally, interpolation between (2.20) and (2.21) yields (2.18). □

Remark 2.6. The estimate in (2.18) does not provide an optimal rate, in fact it is the worst case scenario, as further smoothing may come from $\|b\|_{L^1(0,t)}$. The rate can be improved if we impose further regularity assumptions on b. Indeed, if in addition, b satisfies Assumption 2 from Subsection 4.2, then by (4.3) and (4.8) it follows that $\hat{b}(\lambda) \sim \lambda^{1-\rho}$ as $\lambda \to \infty$. Thus, the nonnegativity of b implies that $\|b\|_{L^1(0,t)} \le C t^{\rho-1}$ by a Tauberian theorem for the Laplace transform (see, for example, [22, Chapter V, Theorem 4.3]). Therefore, in this case, we get a sharper estimate

$$\|A^{-s'/2} \dot{S}(t)\|_{\mathcal{B}(H)} \le C_1 \, t^{\rho s'/2 - 1}, \quad t > 0, \ 0 \le s' \le 2.$$

Nevertheless, the rate in given (2.18) is sufficient for our needs when it is used in the deterministic error analysis for smooth initial data.

2.3. **The continuous stochastic problem.** Next we recall an existence result for the problem (2.1) and, for the sake of completeness, we indicate a proof (see [3, Theorem 2.1] and we refer to [18] for more general noise).

Proposition 2.7. *Let A and Q satisfy* (2.8)–(2.9) *and let b satisfy Assumption 1. Then there exists an unique H-valued (Gaussian) weak solution u of* (2.1) *given by the variation of constants formula*

$$u(t) = S(t)u_0 + \int_0^t S(t-s) \, dW^Q(s). \tag{2.23}$$

Furthermore, the stochastic convolution term has a version whose trajectories are a.s. θ-Hölder continuous for any $\theta < (1 - \rho(\alpha - \kappa))/2$.

Proof. Analogously to [4, Theorem 5.4], it is sufficient to show that the stochastic convolution is well-defined. By Itô's Isometry,

$$\begin{aligned}
\mathbb{E}\left\|\int_0^t S(t-s)\,dW^Q(s)\right\|^2 &= \int_0^t \|S(t-s)Q^{1/2}\|^2_{\mathcal{L}_2(H)} ds \\
&= \int_0^t \sum_{i\geq 1} \|S(t-s)Q^{1/2}e_i\|^2\,ds \\
&= \int_0^t \sum_{i,j\geq 1} (S(t-s)Q^{1/2}e_i, e_j)^2\,ds \\
&= \sum_{j\geq 1}\sum_{i\geq 1}\left(\int_0^t s_j^2(t-s)\,ds\right)(Q^{1/2}e_i,e_j)^2\,ds \\
&\leq C_0 \sum_{j\geq 1}\sum_{i\geq 1} \lambda_j^{-1/\rho}(Q^{1/2}e_i,e_j)^2 \\
&= C_0\|A^{-1/(2\rho)}Q^{1/2}\|^2_{\mathcal{L}_2(H)},
\end{aligned}$$

where we have used Parseval's identity, (2.13) and (2.16). By (2.9) we have that $-1/\rho-\kappa<-\alpha$, and thus using also (2.4),

$$\begin{aligned}
\|A^{-1/(2\rho)}Q^{1/2}\|^2_{\mathcal{L}_2(H)} &= \mathrm{Tr}(A^{-1/\rho}Q) = \mathrm{Tr}(A^{-1/\rho-\kappa}A^{\kappa}Q) \\
&\leq \mathrm{Tr}(A^{-1/\rho-\kappa})\|A^{\kappa}Q\|^2_{\mathcal{B}(H)} \leq \mathrm{Tr}(A^{-\alpha})\|A^{\kappa}Q\|_{\mathcal{B}(H)}.
\end{aligned}$$

Finally, the proof of the Hölder regularity in time of u uses similar techniques and is omitted. □

Remark 2.8. Note that assumptions (2.8)–(2.9) are stronger than the minimal assumption $\|A^{-1/(2\rho)}Q^{1/2}\|_{\mathcal{L}^2(H)}<+\infty$ needed for the existence of a mean squared continuous solution. One can replace (2.8)–(2.9) by

$$\|A^{(s-\frac{1}{\rho})/2}Q^{1/2}\|_{\mathcal{L}^2(H)}<+\infty$$

for some $s>0$ as a single main assumption on A and Q and obtain Hölder regularity of order less than $\min(\frac{1}{2},\frac{\rho s}{2})$.

3. Space discretization

In this section we discretize (2.1) in space by a standard piecewise continuous finite element method. We refer to the monograph [19] for further details on finite elements. We shall derive strong error estimates for the spatially semidiscrete problem for smooth initial data only imposing Assumption 1 on b. We will see later that for time discretization and also for the fully discrete scheme, we have to put further restrictions on b. Let $\{\mathcal{T}_h\}_{0<h<1}$ denote a family of triangulations of $\mathcal{D}$, with mesh size $h>0$ and consider finite element spaces $\{V_h\}_{0<h<1}$, where $V_h\subset H_0^1(\mathcal{D})$ consists of continuous piecewise linear functions vanishing at the boundary of $\mathcal{D}$. In order to derive the finite element formulation we look for a V_h-valued process u_h such that

$$\begin{cases}
(\mathrm{d}u_h(t),\chi) + \displaystyle\int_0^t b(t-s)(\nabla u_h(t),\nabla\chi)\,\mathrm{d}s\,\mathrm{d}t = (\mathrm{d}W^Q(t),\chi),\ \chi\in V_h,\ t>0,\\
(u_h(0),\chi) = (u_0,\chi).
\end{cases}$$

We introduce the "discrete Laplacian"

$$A_h : V_h \to V_h, \quad (A_h\psi, \chi) = (\nabla\psi, \nabla\chi), \quad \psi, \chi \in V_h, \tag{3.1}$$

and the orthogonal projector

$$P_h : H \to V_h, \quad (P_h f, \chi) = (f, \chi), \quad \chi \in V_h.$$

It is clear that the operator A_h is a positive definite bounded operator on V_h. Let us note also that using the definition (3.1) of A_h, the following uniform inequality can be easily derived

$$\|A_h^{-1/2} P_h x\| \le \|A^{-1/2} x\|, \quad x \in H. \tag{3.2}$$

Then, using the L_2-stability of P_h and some interpolation theory, we also have that

$$\|A_h^{-\delta} P_h x\| \le \|A^{-\delta} x\|, \quad \delta \in [0, \frac{1}{2}], \quad x \in H. \tag{3.3}$$

Similarly to $-A$, the operator $-A_h$ generates an analytic contraction semigroup on V_h and satisfies the uniform resolvent estimate

$$\|z(z + A_h)^{-1} P_h\| = \|zR(z, A_h) P_h\| \le M_\phi,$$

for $z \in \Sigma_\phi = \{z \in \mathbb{C} : |\arg(z)| < \phi < \pi\}$. Since $A_h R(z, A_h) = I - zR(z, A_h)$, it follows that

$$\|A_h R(z, A_h) P_h\|_{\mathcal{B}(H)} \le M_\phi + 1, \quad z \in \Sigma_\phi. \tag{3.4}$$

Then we can rewrite the spatially semidiscrete problem in the same form as the original one as

$$\begin{cases} \mathrm{d}u_h + \left(\int_0^t b(t-s) A_h u_h(s) \,\mathrm{d}s \right) \mathrm{d}t = P_h \,\mathrm{d}W^Q(t), \quad t > 0, \\ u_h(0) = P_h u_0. \end{cases} \tag{3.5}$$

Similarly to the original problem the weak solution is given by

$$u_h(t) = S_h(t) P_h u_0 + \int_0^t S_h(t-s) P_h \,\mathrm{d}W^Q(s),$$

where the resolvent family $\{S_h(t)\}_{t \ge 0}$ can be written explicitly as

$$S_h(t) P_h u_0 = \sum_{k=1}^{\infty} s_{h,k}(t) (u_0, e_{h,k}) e_{h,k}.$$

Here $(\lambda_{h,k}, e_{h,k})$ are the eigenpairs of A_h and $s_{h,k}(t)$ are the solution of the ODEs

$$\dot{s}_{h,k}(t) + \lambda_{h,k} \int_0^t b(t-s) s_{h,k}(s) \,\mathrm{d}s = 0, \ s_{h,k}(0) = 1.$$

We have the following stability result.

Lemma 3.1. *If b satisfies Assumption 1, then for some $C > 0$,*

$$\int_0^t \|S(s)x\|^2 \,\mathrm{d}s \le C \|x\|^2_{-\frac{1}{\rho}}, \ t > 0,$$

and

$$\int_0^t \|S_h(s) P_h x\|^2 \,\mathrm{d}s \le C \|x\|^2_{-\frac{1}{\rho}}, \ t > 0, \ h > 0.$$

Proof. We have, by (2.13) and (2.16), that

$$\int_0^t \|S(s)x\|^2 \,\mathrm{d}s = \sum_{k=1}^{\infty} \int_0^t s_k^2(s)\,\mathrm{d}s\,(x, e_k)^2$$

$$\leq \sum_{k=1}^{\infty} \|s_k\|_{L^\infty(\mathbb{R}_+)} \|s_k\|_{L^1(\mathbb{R}_+)} (x, e_k)^2 \leq C_0 \sum_{k=1}^{\infty} \lambda_k^{-1/\rho}(x, e_k)^2 = C_0\|x\|^2_{-\frac{1}{\rho}}.$$

As the constants in (2.13) and (2.16) do not depend on λ_k, we similarly obtain

$$\int_0^t \|S_h(s)P_h x\|^2 \,\mathrm{d}s \leq C_0 \|A_h^{-1/2\rho} P_h x\|^2.$$

Finally, since $-1/2 < -1/2\rho < -1/4$, using (3.3) with $\delta = 1/(2\rho)$, completes the proof. □

The error analysis is based on the Ritz projection

$$R_h : H_0^1(\mathcal{D}) \to V_h, \quad (\nabla R_h v, \nabla \chi) = (\nabla v, \nabla \chi),\ v \in H_0^1(\mathcal{D}),\ \chi \in V_h.$$

In particular, we assume that R_h satisfies the error bound

$$\|R_h v - v\| \leq Ch^\gamma \|v\|_\gamma,\ v \in D(A^{\gamma/2}),\ 1 \leq \gamma \leq 2. \tag{3.6}$$

This puts some restriction on the domain $\mathcal{D}$ but it is satisfied for convex polygonal domains, for instance.

Next we prove an $L^2((0,\infty), H)$ error estimate for the space semidiscretization of the deterministic problem. It is an extension of the result in [2] where the special kernel $b(t) = \frac{1}{\Gamma(\beta)} e^{-t} t^{\beta-1}$ was considered.

Proposition 3.2. *If b satisfies Assumption 1 and* (3.6) *holds, then*

$$\int_0^\infty \|S(t)x - S_h(t)P_h x\|^2 \,\mathrm{d}t \leq Ch^{2s}\|x\|^2_{s-\frac{1}{\rho}},\ 0 \leq s \leq 2.$$

Proof. It follows from Lemma 3.1 that

$$\int_0^t \|(S(s) - S_h(s)P_h)x\|^2 \,\mathrm{d}s \leq 2\int_0^t \|S(s)x\|^2 + \|S_h(s)P_h x\|^2 \,\mathrm{d}s \leq C\|x\|^2_{-\frac{1}{\rho}}. \tag{3.7}$$

To prove an error estimate of optimal order we set

$$\begin{aligned} e(t) &:= S(t)x - S_h(t)P_h x := v(t) - v_h(t) \\ &= v(t) - P_h v(t) + P_h v(t) - v_h(t) := \rho(t) + \theta(t). \end{aligned}$$

For ρ, using the best approximation property of P_h, we obtain by Lemma 3.1 and (3.6) that

$$\int_0^\infty \|\rho(t)\|^2 \,\mathrm{d}t \leq \int_0^\infty \|(R_h - I)v(t)\|^2 \,\mathrm{d}t \leq Ch^4\|x\|^2_{2-\frac{1}{\rho}}. \tag{3.8}$$

In a standard way one derives an equation for θ which reads

$$\begin{cases} \dot{\theta}(t) + \displaystyle\int_0^t b(t-s)A_h\theta(s)\,\mathrm{d}s = A_h P_h \int_0^t b(t-s)(R_h - I)v(s)\,\mathrm{d}s,\ t > 0, \\ \theta(0) = 0. \end{cases}$$

Taking Laplace transforms of both sides yields

$$z\widehat{\theta}(z) + \widehat{b}(z)A_h\widehat{\theta}(z) = A_h P_h (R_h - I)\widehat{v}(z)\widehat{b}(z).$$

Therefore,

$$\widehat{\theta}(z) = A_h R(\frac{z}{\widehat{b}(z)}, A_h) P_h (R_h - I)\widehat{v}(z). \tag{3.9}$$

It can be shown that $\widehat{b}$ extends continuously to $i\mathbb{R} \setminus \{0\}$, see, for example, [14]. Therefore, using (2.6), it follows that $\frac{ik}{\widehat{b}(ik)} \in \Sigma_\phi$, $k \in \mathbb{R} \setminus \{0\}$, with $\phi < \pi$. Thus, $\|A_h R(\frac{ik}{\widehat{b}(ik)}, A_h) P_h\|_{\mathcal{B}(H)} \le (M_\phi + 1)$ by (3.4). Therefore, setting $z = ik$, $k \in \mathbb{R}\setminus\{0\}$, in (3.9) and using the isometry property of the Fourier transform we obtain, by Lemma 3.1 and (3.6), that

$$\int_0^\infty \|\theta(t)\|^2 \,\mathrm{d}t \le (M_\phi + 1) \int_0^\infty \|(R_h - I)v(t)\|^2 \,\mathrm{d}t \le Ch^4 \|x\|^2_{2-\frac{1}{\rho}}. \tag{3.10}$$

Interpolation using (3.7), (3.8), and (3.10) yields

$$\int_0^\infty \|e(t)\|^2 \,\mathrm{d}t \le 2 \int_0^\infty \left(\|\rho(t)\|^2 + \|\theta(t)\|^2\right) \,\mathrm{d}t \le Ch^{2s} \|x\|^2_{s-\frac{1}{\rho}},\ 0 \le s \le 2.$$

□

Next, using the error analysis from [13], we have the following pointwise smooth data estimate for the spatially semidiscrete scheme.

Proposition 3.3. *If b satisfies Assumption 1 and* (3.6) *holds, then for every $\epsilon > 0$ there is $C = C(T, \epsilon)$ such that*

$$\|S(t)x - S_h(t)P_h x\| \le Ch^s \|x\|_{s(1+\epsilon)},\ 0 \le s \le 2,\ t \in [0, T].$$

Proof. As already observed, Assumption 1 implies that b is a positive definite kernel. Therefore by, [13, Theorem 2.1], it follows that

$$\|S(t)x - S_h(t)P_h x\| \le Ch^2 \left(\|x\|_2 + \int_0^t \|\dot{S}(s)x\|_2 \,\mathrm{d}s \right).$$

Proposition 2.5 implies that

$$\int_0^t \|\dot{S}(s)x\|_2 \,\mathrm{d}s = \int_0^t \|A^{-\epsilon}\dot{S}(s)A^{1+\epsilon}x\| \,\mathrm{d}s \le C(T, \epsilon)\|x\|_{2+2\epsilon}. \tag{3.11}$$

Finally, since $\|S(t) - S_h(t)P_h\|_{\mathcal{B}(H)} \le 2$, interpolation finishes the proof. □

Theorem 3.4. *Let A and Q satisfy (2.8)–(2.9) and let b be satisfy Assumption 1. If $\mathbb{E}\|u_0\|^2_{\nu(1+\epsilon)} < \infty$ and* (3.6) *holds, then there is $C = C(T, \epsilon, \nu)$ such that*

$$\left(\mathbb{E}\|u(t) - u_h(t)\|^2\right)^{1/2} \le Ch^\nu,\ \nu \le \frac{1}{\rho} - \alpha + \kappa,\ t \in [0, T].$$

Proof. By the variation of constants formula,

$$u(t) - u_h(t) = S(t)x - S_h(t)x + \int_0^t (S(t-s) - S_h(t-s)P_h) \,\mathrm{d}W^Q(s).$$

Thus,

$$\begin{aligned}\mathbb{E}\|u(t) - u_h(t)\|^2 &\le 2\mathbb{E}\|S(t)x - S_h(t)x\|^2 \\ &\quad + 2\mathbb{E}\left\| \int_0^t (S(t-s) - S_h(t-s)P_h) \,\mathrm{d}W^Q(s) \right\|^2 := e_1 + e_2.\end{aligned}$$

It follows from Proposition 3.3 that

$$e_1 \le Ch^{2\nu}\mathbb{E}\|u_0\|^2_{\nu(1+\epsilon)}.$$

To bound e_2 we use Itô's Isometry and Proposition 3.2 to obtain

$$\begin{aligned}
e_2 &= \mathbb{E}\left\|\int_0^t (S(t-s) - S_h(t-s)P_h)\,\mathrm{d}W^Q(s)\right\|^2 \\
&= \int_0^t \|(S(t-s) - S_h(t-s)P_h)Q^{1/2}\|^2_{\mathcal{L}_2(H)}\,\mathrm{d}s \\
&= \sum_{k=1}^{\infty}\int_0^t \|(S(s) - S_h(s)P_h)Q^{1/2}e_k\|^2\,\mathrm{d}s \\
&\le Ch^{2\nu}\sum_{k=1}^{\infty}\|A^{(\nu-\frac{1}{\rho})/2}Q^{1/2}e_k\|^2 = Ch^{2\nu}\|A^{(\nu-\frac{1}{\rho})/2}Q^{1/2}\|^2_{\mathcal{L}_2(H)} \\
&= Ch^{2\nu}\mathrm{Tr}(A^{\nu-\frac{1}{\rho}}Q) \le Ch^{2\nu}\mathrm{Tr}(A^{\nu-\frac{1}{\rho}-\kappa})\|A^{\kappa}Q\|.
\end{aligned} \tag{3.12}$$

□

Remark 3.5. In particular, if $Q = I$, then $d = 1$, $\kappa = 0$ and $\alpha > \frac{1}{2}$ whence $\nu < \frac{1}{\rho} - \frac{1}{2}$. Also note, that it is clear from the proof that instead of (2.8)–(2.9) we could assume that $\|A^{(\nu-\frac{1}{\rho})/2}Q^{1/2}\|_{\mathcal{L}_2(H)} < \infty$ and get a convergence rate of order ν. Then, for trace class noise; that is, when $\mathrm{Tr}(Q) < \infty$ we can take $\nu = \frac{1}{\rho}$.

We end this section by showing that the above error estimate is optimal in the sense that it corresponds to the spatial regularity of the solution.

Theorem 3.6. *Let A and Q satisfy (2.8)–(2.9) and let $\nu = \frac{1}{\rho} - \alpha + \kappa$, or, let $\|A^{(\nu-\frac{1}{\rho})/2}Q^{1/2}\|_{\mathcal{L}_2(H)} < \infty$ for some $\nu \ge 0$. If b satisfies Assumption 1 and $\mathbb{E}\|u_0\|^2_\nu < \infty$, then $\mathbb{E}\|u(t)\|^2_\nu \le C$ for some $C > 0$ for all $t \ge 0$.*

Proof. It follows by Itô's Isometry and the fact that $\|S(t)\| \le 1$ that

$$\mathbb{E}\|u(t)\|^2_\nu \le 2\mathbb{E}\|u_0\|^2_\nu + 2\int_0^t \|A^{\nu/2}S(s)Q^{1/2}\|^2_{\mathcal{L}_2(H)}\mathrm{d}s.$$

Let (e_k, λ_k) be the eigenpairs of A. Then, by monotone convergence, the self-adjointness of A and S, and Proposition 2.4, it follows that

$$\begin{aligned}
&\int_0^t \|A^{\nu/2}S(s)Q^{1/2}\|^2_{\mathcal{L}_2(H)}\mathrm{d}s = \sum_{k=1}^{\infty}\int_0^t \|A^{\nu/2}S(s)Q^{1/2}e_k\|^2\,\mathrm{d}s \\
&= \sum_{j,k=1}^{\infty}\int_0^t (A^{\nu/2}S(s)Q^{1/2}e_k, e_j)^2\,\mathrm{d}s = \sum_{j,k=1}^{\infty}\int_0^t (Q^{1/2}e_k, S(s)A^{\nu/2}e_j)^2\,\mathrm{d}s \\
&= \sum_{j,k=1}^{\infty}(Q^{1/2}e_k, \lambda_j^{n/2}e_j)^2\int_0^t s_j^2(s)\,\mathrm{d}s \\
&\le \sum_{j,k=1}^{\infty}(Q^{1/2}e_k, \lambda_j^{\nu/2}e_j)^2\|s_j\|_{L^\infty(\mathbb{R}_+)}\|s_j\|_{L^1(\mathbb{R}_+)} \\
&\le C_0\sum_{j,k=1}^{\infty}(Q^{1/2}e_k, \lambda_j^{\nu/2}e_j)^2\lambda_j^{-1/\rho} = C_0\sum_{j,k=1}^{\infty}(Q^{1/2}e_k, \lambda_j^{\nu/2-\frac{1}{2\rho}}e_j)^2 \\
&= C_0\|A^{(\nu-\frac{1}{\rho})/2}Q^{1/2}\|^2_{\mathcal{L}_2(H)} \le C_0\mathrm{Tr}(A^{\nu-\frac{1}{\rho}-\kappa})\|A^\kappa Q\|^2_{\mathcal{B}(H)}.
\end{aligned}$$

□

4. Time discretization

Time discretization is achieved via a classical implicit Euler scheme and, concerning the convolution in time, via a quadrature rule based on (1.7). Let $\Delta t > 0$ and we set $t_n = n\,\Delta t$ for any integer $n \ge 0$. We seek for an approximation u_n of $u(t_n)$ defined by the recurrence

$$u_n - u_{n-1} + \Delta t\left(\sum_{k=1}^{n}\omega_{n-k}\,Au_k\right) = W^Q(t_n) - W^Q(t_{n-1}), \quad n \ge 1, \tag{4.1}$$

with initial condition $u_0 = u(0)$. We recall that the coefficients $\{\omega_k\}_{k\ge 0}$ of the quadrature are chosen such that

$$\sum_{k=0}^{+\infty}\omega_k z^k = \widehat{b}\left(\frac{1-z}{\Delta t}\right), \quad |z| < 1. \tag{4.2}$$

Let us note that thanks to [14, estimate (3.6)], we have the lower bound for ω_0:

$$\omega_0 = \widehat{b}(1/\Delta t) \ge c\Delta t^{\rho-1},\ \Delta t < 1, \tag{4.3}$$

where $\rho \in (1, 2)$ is defined in (2.6).

In the sequel we derive a discrete mild formulation (variation of constants formula) for (4.1). This formulation can not be made easily explicit as a function of the operators A, Q and the kernel b, because of the memory effect in the drift. First consider the deterministic algorithm

$$v_n - v_{n-1} + \Delta t\left(\sum_{k=1}^{n}\omega_{n-k}\,Av_k\right) = 0, \quad n \ge 1; \quad v_0 = x. \tag{4.4}$$

Taking the z-transform, using the notation

$$\hat{V}(z) = \sum_{k=0}^{\infty}v_k z^k \text{ and } \hat{\omega}(z) = \sum_{k=0}^{\infty}\omega_k z^k,$$

we get

$$\hat{V}(z) - x - z\hat{V}(z) + \Delta t\hat{\omega}(z)A(\hat{V}(z) - x) = 0.$$

Thus,

$$\hat{V}(z) = (I + \Delta t\hat{\omega}(z)A)((1-z)I + \Delta t\hat{\omega}(z)A)^{-1}x := \hat{B}(z)x,$$

where

$$\hat{B}(z)x = \sum_{k=0}^{\infty} B_k x z^k.$$

This means that $v_k = B_k x$, $k = 0, 1, ...$ Note that $B_0 = \hat{B}(0) = I$. For the stochastic equation it will be useful to rewrite $\hat{B}(z)x$ as

(4.5)
$$\begin{aligned}\hat{B}(z)x =&((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}(I + \hat{\omega}(z)\Delta tA)x \\ &= ((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}x + \hat{\omega}(z)\Delta tA((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}x \\ &= ((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}x - (1-z)((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}x + x \\ &= (z((1-z)I + \hat{\omega}(z)\Delta tA)^{-1} + I)x.\end{aligned}$$

Now, we consider the stochastic case (4.1) which reads, after taking the z-transform, rearranging, and using the notation $w_n = W^Q(t_n) - W^Q(t_{n-1})$ for $n \geq 1$, $w_0 = 0$, and

$$\hat{w}(z) = \sum_{k=0}^{\infty} w_k z^n \text{ and } \hat{U}(z) = \sum_{k=0}^{\infty} u_k z^k,$$

as

$$\begin{aligned}\hat{U}(z) &= \hat{B}(z)x + ((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}\hat{w}(z) \\ &= \hat{B}(z)x + \frac{\hat{B}(z) - I}{z}\hat{w}(z) = \hat{B}(z)x + \hat{B}(z)\frac{\hat{w}(z)}{z} - \frac{1}{z}\hat{w}(z),\end{aligned}$$

where we also used (4.5) to rewrite the stochastic term in the previous calculation. This yields the discrete variation of constants formula, taking into account that $w_0 = 0$ and that $B_0 = I$,

$$u_n = B_n x + \sum_{k=0}^{n} B_{n-k}w_{k+1} - w_{n+1} = B_n x + \sum_{k=0}^{n-1} B_{n-k}w_{k+1}. \tag{4.6}$$

The importance of this formula lies in the fact that it connects the deterministic case to the stochastic case with the deterministic time-discrete solution operators B_n explicitly appearing in the formula.

4.1. **Deterministic estimates: stability and smoothing.** The next theorem is interesting in its own right. It shows that Lubich's convolution quadrature based on the Backward Euler scheme have a remarkable qualitative property: it preserves the L^p-norm of the orbits of the solution. The result can be viewed as a generalization of the ones in [6]; in particular, it removes the additional technical frequency condition in [6, Theorem 2]. The proof uses a representation similar to that in [1]. We also note that the statement holds in Banach spaces as well since the proof does not use Hilbert space techniques.

Theorem 4.1. *If the resolvent family* $\{S(t)\}_{t\geq 0}$ *of* (2.10) *satisfies*

$$S(\cdot)x \in L^p((0,\infty);H)$$

for some $1 \leq p \leq \infty$ *and* $x \in H$*, then*

$$\Delta t \sum_{k=1}^{n} \|B_k x\|^p \leq \int_0^\infty \|S(t)x\|^p \,\mathrm{d}t, \quad 1 \leq p < \infty,$$

and

$$\sup_{k\geq 1} \|B_k x\| \leq \|S(\cdot)x\|_{L^\infty(\mathbb{R}_+)}.$$

Proof. The Laplace Transform of $\{S(t)\}_{t\geq 0}$ is given by

$$\hat{S}(z)x = (zI + \hat{b}(z)A)^{-1}x.$$

Using (4.2) and (4.5) we see that the z-transform $\hat{B}x$ of $\{B_n x\}_n$ is given by

$$\begin{aligned}\hat{B}(z) &= z\frac{1}{\Delta t}\hat{S}(\frac{1-z}{\Delta t})x + x = x + z\int_0^\infty S(\Delta t s)e^{-s}e^{zs}\,\mathrm{d}s \\ &= x + \sum_{k=1}^{\infty} z^k \int_0^\infty S(\Delta t s)x\frac{e^{-s}s^{k-1}}{(k-1)!}\,\mathrm{d}s.\end{aligned}$$

Therefore, we conclude that $B_0 = I$ and that

$$B_k x = \int_0^\infty S(\Delta t s)x\frac{e^{-s}s^{k-1}}{(k-1)!}\,\mathrm{d}s \text{ for } k \geq 1. \tag{4.7}$$

Let

$$f_k(s) := \frac{e^{-s}s^{k-1}}{(k-1)!}, \quad k \geq 1.$$

Then $f_k \geq 0$, $\|f_k\|_{L^1(\mathbb{R}_+)} = 1$. Therefore, if $p = \infty$, we immediately obtain from (4.7) that

$$\sup_{k\geq 1} \|B_k x\| \leq \|S(\cdot)x\|_{L^\infty(\mathbb{R}_+)}.$$

If $1 \leq p < \infty$, then we use Jensen's inequality in (4.7), and have

$$\begin{aligned}\Delta t \sum_{k=1}^{n} \|B_k x\|^p &\leq \sum_{k=1}^{n} \Delta t \int_0^\infty \|S(\Delta t s)x\|^p f_k(s)\,\mathrm{d}s \\ &= \int_0^\infty \|S(t)x\|^p \sum_{k=1}^{n} f_k(\frac{t}{\Delta t})\,\mathrm{d}t \leq \sup_{t>0}\sum_{k=1}^{\infty} f_k(t)\int_0^\infty \|S(t)x\|^p\,\mathrm{d}t.\end{aligned}$$

Finally, noticing that $\sum_{n=1}^{\infty} f_n \equiv 1$ completes the proof. □

Theorem 4.1 has the following important corollary on the smoothing and stability of the time discretization scheme in case b satisfies Assumption 1.

Corollary 4.2. *If* b *satisfies Assumption 1, then, for all* $x \in H$*,*

$$\sup_{k\geq 1} \|B_k x\| \leq \|x\| \text{ and } \Delta t \sum_{k=1}^{n} \|B_k x\|^2 \leq C\|x\|_{-\frac{1}{\rho}}^2,\ n \geq 1.$$

Proof. The statement follows from Theorem 4.1 together with Lemma 3.1 and the fact that $\|S(t)\| \leq 1$ for $t \geq 0$. □

Finally we will need a Hölder type estimate on the resolvent family $\{S(t)\}_{t\geq 0}$.

Lemma 4.3. *If b satisfies Assumption 1, then there is $C = C(T,\gamma) > 0$ such that*

$$\left(\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} \|(S(t_n-s)-S(t_n-t_{k-1}))x\|^2\,\mathrm{d}s\right)^{1/2} \le C\Delta t^{\gamma}\|x\|_{s-\frac{1}{\rho}},\ n\Delta t = T,$$

for all $\gamma < \frac{\rho s}{2}$ where $0 < s \le \frac{1}{\rho}$.

Proof. It follows from (2.17), with $s = \frac{1}{\rho}-\epsilon$, and Lemma 3.1 that there is a constant $C = C(\epsilon, T)$ such that, for $0 < \epsilon \le \frac{1}{\rho}$,

$$\left(\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} \|(S(t_n-s)-S(t_n-t_{k-1}))x\|^2\,\mathrm{d}s\right)^{1/2} \le C\|x\|_{\epsilon-\frac{1}{\rho}},\ n\Delta t = t_n = T.$$

Next, it follows from Proposition 2.4 that

$$\begin{aligned}
&\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} \|(S(t_n-s)-S(t_n-t_{k-1}))x\|^2\,\mathrm{d}s \\
&\qquad = \sum_{i=1}^{\infty}(x,e_i)^2\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k}(s_i(t_n-s)-s_i(t_n-t_{k-1})^2\,\mathrm{d}s \\
&\qquad \le 2\sum_{i=1}^{\infty}(x,e_i)^2\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k}|s_i(t_n-s)-s_i(t_n-t_{k-1})|\,\mathrm{d}s \\
&\qquad \le 2\sum_{i=1}^{\infty}(x,e_i)^2\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k}\int_{t_n-s}^{t_n-t_{k-1}}|\dot{s}_i(t)|\,\mathrm{d}t\,\mathrm{d}s \\
&\qquad \le 2\sum_{i=1}^{\infty}(x,e_i)^2\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k}\int_{t_n-t_k}^{t_n-t_{k-1}}|\dot{s}_i(t)|\,\mathrm{d}t\,\mathrm{d}s \\
&\qquad \le 2\Delta t\|x\|^2\sup_{i\ge 1}\|\dot{s}_i\|_{L^1(\mathbb{R}_+)} \le C\Delta t\|x\|^2.
\end{aligned}$$

Finally, interpolation gives the desired result. □

4.2. **Deterministic estimates: convergence rates.** In order to give an error estimate of optimal order with no initial regularity for the time discretization of the deterministic problem we have to impose another assumption on b. This kind of assumption; that is, the existence of an analytic extension of $\hat{b}$ to a sector beyond the left halfplane, is fairly standard in the existing deterministic literature, see, for example, [5, 9, 11, 12], but it clearly represents a major restriction compared to Assumption 1. We note that this additional assumption is not needed neither for the spatial error estimates with smooth initial data, and hence for the space-semidiscretization of the stochastic equation, nor the for the stability results for the time discretization in the previous subsection.

Assumption 2. *The Laplace transform $\widehat{b}$ of b can be extended to an analytic function in a sector Σ_θ with $\theta > \pi/2$ and $|\widehat{b}^{(k)}(z)| \le C|z|^{1-\rho-k}$, $k = 0,1$, $z \in \Sigma_\theta$.*

Note that Assumption 2 implies that

$$\omega_0 = \widehat{b}(1/\Delta t) \le C\Delta t^{\rho-1},\ \Delta t < 1. \tag{4.8}$$

An important example of a family of kernels satisfying both Assumptions 1 and 2 is given by $b(t) = Ct^{\beta-1}e^{-\eta t}$, $0 < \beta < 1$ and $\eta \ge 0$.

Assumptions 1 and 2 allows us to use the following deterministic nonsmooth data estimate [12, Theorem 3.2].

Proposition 4.4. *If Assumptions 1 and 2 hold, then there exists $C = C(\rho) > 0$ such that*

$$\|S(t_n)x - B_n x\| \le \frac{C}{t_n}\Delta t\,\|x\|, \quad n \ge 1. \tag{4.9}$$

Corollary 4.5. *If Assumptions 1 and 2 hold, then there exists $C = C(T, \gamma, \rho)$ such that*

$$\left(\Delta t \sum_{k=0}^{n} \|S(t_k)x - B_k x\|^2\right)^{1/2} \le C\Delta t^{\gamma}\|x\|_{s-\frac{1}{\rho}}, \quad n\Delta t = T,$$

for all $\gamma < \frac{\rho s}{2}$ where $0 < s \le \frac{1}{\rho}$.

Proof. It follows from (2.17), with $s = \frac{1}{\rho} - \epsilon$, and Corollary 4.2 that there is a constant $C = C(\epsilon, T)$ such that, for $0 < \epsilon \le \frac{1}{\rho}$,

$$\left(\Delta t \sum_{k=0}^{n} \|S(t_k)x - B_k x\|^2\right)^{1/2} \le C\|x\|_{\epsilon-\frac{1}{\rho}}, \ n\Delta t = T, \quad \epsilon > 0,$$

where we also used the fact that $B_0 = S(t_0) = I$. Furthermore, since $\|S(t_k) - B_k\| \le 2$ by Corollary 4.2, it follows from (4.9) that

$$\|S(t_k)x - B_k x\| \le C\Delta t^{\frac{1}{2}-\epsilon} t_k^{\epsilon-\frac{1}{2}}\|x\|, \quad k \ge 1,$$

and thus, for some $C = C(\epsilon, T, \rho)$,

$$\left(\Delta t \sum_{k=0}^{n} \|S(t_k)x - B_k x\|^2\right)^{1/2} \le C\Delta t^{\frac{1}{2}-\epsilon}\|x\|.$$

Interpolation finishes the proof. □

4.3. **Error estimate for the stochastic equation.** We can now state and proof the main result of this section.

Theorem 4.6. *Let A and Q satisfy (2.8)–(2.9) and let b satisfy Assumptions 1 and 2. Suppose further that $\mathbb{E}\|u_0\|^2 < \infty$. For $T > 0$, let $\{u(t)\}_{t\in[0,T]}$ be the unique weak solution of (2.1) and let u_n be the solution of the scheme (4.1) with $T = n\Delta t$. Then for any $\gamma < (1 - \rho(\alpha - \kappa))/2$, there is $C = C(\rho, \mathbb{E}\|u_0\|^2) > 0$ and $K = K(T, \alpha, \gamma, \kappa, \rho) > 0$ such that*

$$(\mathbb{E}\,\|u(T) - u_n\|^2)^{1/2} \le CT^{-1}\Delta t + K\Delta t^{\gamma}, \quad t_n = n\Delta t = T. \tag{4.10}$$

Proof. If $e_n = u(T) - u_n = u(t_n) - u_n$, then (2.23) and (4.6) yields

$$e_n = (S(t_n) - B_n)u_0 + \sum_{k=1}^{n}\left[\int_{t_{k-1}}^{t_k} (S(t_n - s) - B_{n-k+1})\,\mathrm{d}W^Q(s)\right].$$

Taking the expectation of the square of the H-norm of e_n leads to, by independence and Itô's isometry:

$$\mathbb{E}\|e_n\|^2 \le 2(a + b), \tag{4.11}$$

where a denotes the deterministic part of the error:

$$a = \mathbb{E}\|(S(t_n) - B_n)u_0\|^2, \tag{4.12}$$

and b the stochastic part:

$$b = \sum_{i=1}^{+\infty} \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \|(S(t_n - s) - B_{n-k+1})Q^{1/2}e_i\|^2 \, \mathrm{d}s.$$

Thanks to (4.9), a can be bounded as

$$a \le \frac{C}{t_n^2} \Delta t^2 \mathbb{E}\|u_0\|^2, \quad n \ge 1. \tag{4.13}$$

We use Corollary 4.5 and Lemma 4.3 to bound b as

$$\begin{aligned} b &\le 2 \sum_{i=1}^{\infty} \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \|(S(t_n - s) - S(t_n - t_{k-1}))Q^{1/2}e_i\|^2 \, \mathrm{d}s \\ &\quad + 2 \sum_{i=1}^{\infty} \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \|(S(t_n - t_{k-1}) - B_{n-k+1})Q^{1/2}e_i\|^2 \, \mathrm{d}s \\ &\le C\Delta t^{2\gamma} \sum_{i=1}^{\infty} \|Q^{1/2}e_i\|_{s-\frac{1}{\rho}} = C\Delta t^{2\gamma} \|A^{(s-\frac{1}{\rho})/2} Q^{1/2}\|^2_{\mathcal{L}_2(H)} \\ &\le C\Delta t^{2\gamma} \mathrm{Tr}(A^{s-\frac{1}{\rho}-\kappa}) \|A^{\kappa} Q\|^2_{\mathcal{B}(H)}. \end{aligned}$$

Finally, we set $-\alpha = s - \frac{1}{\rho} - \kappa$ and conclude that $\gamma < \frac{\rho s}{2} = (1 - \rho(\alpha - \kappa))/2$. □

Remark 4.7. In particular, if $Q = I$ then $d = 1$, $\kappa = 0$ and $\alpha > \frac{1}{2}$ whence $\gamma < 1/2 - \frac{\rho}{4}$. Also note, that it is clear from the proof that instead of (2.8)–(2.9) we could assume that $\|A^{(s-\frac{1}{\rho})/2} Q^{1/2}\|_{\mathcal{L}_2(H)} < \infty$ and obtain $\gamma < \frac{\rho s}{2}$. Then, for trace class noise; that is, when $\mathrm{Tr}(Q) < \infty$ we can take $s = \frac{1}{\rho}$ and hence $\gamma < 1/2$. Remarkably, this is the same rate as for the heat equation [21] independently of the value of ρ.

5. The fully discrete scheme

In this section we derive strong error estimates for a fully discrete scheme for (2.1). Both Assumptions 1 and 2 on b are needed but in return we get optimal error bounds with no initial regularity. As the fully discrete scheme, similarly to the time semidiscretization (4.1), we consider the recurrence

$$u_{n,h} - u_{n-1,h} + \Delta t \left(\sum_{k=1}^{n} \omega_{n-k} \, A_h u_{k,h} \right) = P_h(W^Q(t_n) - W^Q(t_{n-1})), \quad n \ge 1, \tag{5.1}$$

with $u_{0,h} = P_h u_0$. Again, the solution is given by the discrete variation of constants formula

$$u_{n,h} = B_{n,h} P_h u_0 + \sum_{k=0}^{n-1} B_{n-k,h} P_h \Delta W^Q_{k+1}, \tag{5.2}$$

where $\Delta W_{k+1} = W(t_{k+1}) - W(t_k)$ and $\{B_{k,h}\}_{k \ge 0}$ is a family of linear bounded operators with $B_{0,h} = I$.

Theorem 5.1. *Let A and Q satisfy (2.8)–(2.9) and let b be satisfy Assumptions 1 and 2. Suppose further that $\mathbb{E}\|u_0\|^2 < \infty$. For $T > 0$, let $\{u(t)\}_{t \in [0,T]}$ be the unique weak solution of* (2.1) *and let $u_{n,h}$ be the solution of the scheme* (5.1) *with $T = n\Delta t$.*

If (3.6) *holds, then there is* $C = C(\rho, \mathbb{E}\|u_0\|^2) > 0$ *and* $K = K(T, \alpha, \gamma, \kappa, \rho) > 0$ *such that*

$$(5.3)\qquad \left(\mathbb{E}\|u(T) - u_{n,h}\|^2\right)^{1/2} \le C(\Delta t T^{-1} + h^2 T^{-\rho}) + K(\Delta t^{\gamma} + h^{\nu}),\ n\Delta t = T,$$

where $\gamma < (1 - \rho(\alpha - \kappa))/2$ *and* $\nu \le \frac{1}{\rho} - \alpha + \kappa$.

Proof. We decompose the error as

$$\begin{aligned} u(T) - u_{n,h} &= S(T)u_0 - B_{n,h}P_h u_0 \\ &\quad + \int_0^T S(T-s)\,\mathrm{d}W^Q(s) - \int_0^T S_h(T-s)P_h\,\mathrm{d}W^Q(s) \\ &\quad + \int_0^T S_h(T-s)P_h\,\mathrm{d}W^Q(s) - \sum_{k=0}^{n-1} B_{n-k,h}P_h \Delta W^Q_{k+1} \\ &:= e_1 + e_2 + e_3. \end{aligned}$$

First we bound e_1 which is the deterministic error. Under Assumptions 1 and 2 we have that

$$(\mathbb{E}\|e_1\|^2)^{1/2} \le C(\Delta t T^{-1} + h^2 T^{-\beta-1})(\mathbb{E}\|u_0\|^2)^{1/2}$$

by [12, Theorems 2.1 and 3.2]. Next, e_2 has already been bounded in (3.12) as

$$(5.4)\qquad \mathbb{E}\|e_2\|^2 \le Ch^{2\nu}\|A^{(\nu-\frac{1}{\rho})/2}Q^{1/2}\|^2_{\mathcal{L}_2(H)} \le Ch^{2\nu}\mathrm{Tr}(A^{\nu-\frac{1}{\rho}-\kappa})\|A^{\kappa}Q\|.$$

Finally, the proof of Theorem 4.6 shows that,

$$(5.5)\qquad \mathbb{E}\|e_3\|^2 \le C\Delta t^{2\gamma}\|A_h^{(s-\frac{1}{\rho})/2}(P_hQP_h)^{1/2}\|^2_{\mathcal{L}_2(H)}.$$

Set $-r = (s - \frac{1}{\rho})/2$ and note that since $0 < s \le \frac{1}{\rho}$ we have that $0 \le r < 1/2$. Then,

$$\begin{aligned} \|A_h^{-r}(P_hQP_h)^{1/2}\|^2_{\mathcal{L}_2(H)} &= \mathrm{Tr}(P_hA_h^{-r}P_hQP_hA_h^{-r}P_h) = \|A_h^{-r}P_hQ^{1/2}\|^2_{\mathcal{L}_2(H)} \\ &\le \|A_h^{-r}P_hA^{r}\|^2_{\mathcal{B}(H)}\|A^{-r}Q^{1/2}\|^2_{\mathcal{L}_2(H)}. \end{aligned}$$

Thanks to (3.3) with $\delta = r \in [0, 1/2)$, it follows that $\|A_h^{-r}P_hA^r\|_{\mathcal{B}(H)} \le 1$. Hence,

$$\mathbb{E}\|e_3\|^2 \le C\Delta t^{2\gamma}\|A^{(s-\frac{1}{\rho})/2}Q^{1/2}\|^2_{\mathcal{L}_2(H)} \le C\Delta t^{2\gamma}\mathrm{Tr}(A^{s-\frac{1}{\rho}-\kappa})\|A^{\kappa}Q\|^2_{\mathcal{B}(H)},$$

and the proof is complete. □

Remark 5.2. We would like to highlight two important special cases. Firstly, if $Q = I$ then $d = 1$, $\kappa = 0$ and $\alpha > \frac{1}{2}$. Hence $\nu < \frac{1}{\rho} - \frac{1}{2}$ and $\gamma < 1/2 - \frac{\rho}{4}$. As before, we could assume, that $\|A^{(\nu-\frac{1}{\rho})/2}Q^{1/2}\|_{\mathcal{L}_2(H)} < \infty$ instead of (2.8) and (2.9) and get a convergence rate of order ν is space and $\gamma < \frac{\rho\nu}{2}$ in time. In particular, if $\mathrm{Tr}(Q) < \infty$, then we may set $\nu = \frac{1}{\rho}$. Thus, the time order is almost 1/2, the same as for the heat equation with trace class noise, but the space order is less than 1, which is the space order for the heat equation, see [21].

Remark 5.3. The pure time-discretization as well as the fully discrete scheme can be studied for smooth initial data under Assumptions 1 and 2 on b. Using [13, Theorem 3.1] and [12, Lemma 3.2] one arrives at the deterministic estimate

$$(5.6)\qquad \begin{aligned} &\|S(T)u_0 - B_{n,h}P_hu_0\| \\ &\quad \le C(T)(h^2 + k)\left(\|u_0\|_2 + \int_0^T \|\dot{S}(s)u_0\|_2\,\mathrm{d}s + \int_0^T \|\ddot{S}(s)u_0\|\,\mathrm{d}s\right). \end{aligned}$$

If $u_0 \in \mathcal{D}(A)$, then $u(t) = S(t)u_0$ is a strong solution of (2.10), see [17, Proposition 1.2]; that is, $u(t) = S(t)u_0$ satisfies (2.10) with $f \equiv 0$ for all $t > 0$. Then

$$\ddot{S}(t)u_0 + \int_0^t b(t-s)A\dot{S}(s)u_0 \,\mathrm{d}s + b(t)Au_0 = 0,\ t > 0,$$

and thus

$$\int_0^T \|\ddot{S}(s)u_0\| \,\mathrm{d}s \le C(T) \left(\int_0^T \|\dot{S}(s)u_0\|_2 \,\mathrm{d}s + \|u_0\|_2 \right).$$

Therefore, using stability, interpolation and (3.11), it follows that

$$\|S(T)u_0 - B_{n,h}P_h u_0\| \le C(T,\epsilon)(h^s + k^{s/2})\|u_0\|_{s(1+\epsilon)},\ 0 \le s \le 2.$$

The latter estimate can be used to replace the first term in the bound (5.3) in case $\mathbb{E}\|u_0\|^2_{s(1+\epsilon)} < \infty$, $0 \le s \le 2$. The estimates for the pure time-discretization are analogous using [12, Theorem 3.1] which states (5.6) with $h = 0$ and $B_{n,0} = B_n$.

References

[1] M. P. Calvo, E. Cuesta and C. Palencia, Runge-Kutta convolution quadrature methods for well-posed equations with memory, Numer. Math., 107(4), 589–614 (2007).

[2] U. J. Choi and R. C. Maccamy, Fractional order Volterra equations, Volterra integrodifferential equations in Banach spaces and applications (Trento, 1987), Pitman Res. Notes Math. Ser., 190, Longman Sci. Tech., Harlow, 1989, 231–245.

[3] P. Clément, G. Da Prato and J. Prüss, White noise perturbation of the Equations of Linear Parabolic Viscoelasticity, Rend. Istit. Mat. Univ. Trieste, XXIX, 207–220 (1997).

[4] G. Da Prato and J. Zabczyk, Stochastic Equations in Infinite Dimensions, xviii+454 pp. Cambridge University Press, Cambridge (1992).

[5] P. P. B. Eggermont, On the quadrature error in operational quadrature methods for convolutions, Numer. Math., 62, 35–48 (1992).

[6] C. B. Harris and R. D. Noren, Uniform l^1 behavior of a time discretization method for a Volterra integrodifferential equation with convex kernel; stability, SIAM J. Numer. Anal., 49(4), 1553–1571 (2011).

[7] A. Karczewska and P. Rozmej, On Numerical Solutions to stochastic Volterra equations, arXiv:math/0409026.

[8] M. Kovács, S. Larsson and F. Lindgren, Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise, Numer. Algorithms, 53(2), 309–320 (2010).

[9] C. Lubich, Convolution quadrature and discretized operational calculus. I, Numer. Math., 52(2), 129–145 (1988).

[10] C. Lubich, Convolution quadrature and discretized operational calculus. II, Numer. Math.. 52(4), 413–425 (1988).

[11] C. Lubich, Convolution quadrature revisited, B.I.T., 44(3), 503–514 (2004).

[12] C. Lubich, I. Sloan and V. Thomée, Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term, Math. Comp. 65, 1–17 (1996).

[13] W. McLean and V. Thomée, Numerical solution of an evolution equation with a positive-type memory term, J. Austral. Math. Soc. Ser. B, 35, 23–70 (1993).

[14] S. Monniaux and J. Prüss, A theorem of the Dore-Venni type for noncommuting operators, Trans. Amer. Math. Soc., 349, 4787–4814 (1997).

[15] J. A. Nohel and D. F. Shea, Frequency domain methods for Volterra equations, Adv. Math., 22, 278–304 (1976).

[16] J. Printems, On the discretization in time of parabolic stochastic partial differential equations, Math. Model. and Numer. Anal., 35(6), 1055–1078 (2001).

[17] J. Prüss, Evolutionary Integral Equations and Applications, xxvi+366 pp. Birkhäuser Verlag, Basel (1993).

[18] S. Sperlich, On parabolic Volterra equations disturbed by fractional Brownian motions, Stoch. Anal. Appl., 27(1), 74–94 (2009).

[19] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, 2nd ed., xii+370 pp. Springer-Verlag, Berlin (2006).
[20] D. Xu, Stability of the difference type methods for linear Volterra equations in Hilbert spaces, Numer. Math., 109(4), 571–595 (2008).
[21] Y. Yan, Galerkin finite element methods for stochastic parabolic partial differential equations, SIAM J. Numer. Anal., 43(4), 1363–1384 (2005).
[22] D. V. Widder, The Laplace Transform, x+406 pp. Princeton University Press, Princeton, N.J. (1941).

Department of Mathematics and Statistics, University of Otago, PO Box 56, Dunedin, 9054, New Zealand.
E-mail address: mkovacs@maths.otago.ac.nz

Laboratoire d'Analyse et de Mathématiques Appliquées, CNRS UMR 8050, 61, avenue du Général de Gaulle, Université Paris–Est, 94010 Créteil, France.
E-mail address: printems@u-pec.fr